DETAILED ACTION

Claim Rejections - 35 USC § 102

1.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

2.   Claims 1 to 2, 4 to 6, and 20 to 24 are rejected under 35 U.S.C. 102(e) as being

anticipated by *Moore et al. ('041)*.

Concerning independent claim 1, *Moore et al. ('041)* discloses a method for

responding to messages, comprising:

"receiving at a speech processing device a speech input received by said

teleconferencing system from a telephone connected to the teleconferencing system" –

speech information from a caller using telephone 62 is carried through PSTN 60, and is

directed to speech-to-text module 74 of intelligent media translator 70 ("a speech

processing device") through service provider system 30 ("said teleconferencing

system") via line 80 (Page 10, ¶[0118]: Figure 1); service provider system 30 may

include, without limitation, conference call establishment (Page 7: ¶[0087]); thus,

service provider system 30 is equivalent to "said teleconferencing system"; instant

messaging (IM) service 22 ("the instant messaging system") communicates instant messages through chat client 14 so that text instant messages can be exchanged in real time with one or more parties (Page 7, ¶[0077]- ¶[0082]: Figure 1);

"transcribing the speech into a first text message by the speech processing device" – the packetized data stream is directed to speech-to-text module 74 of intelligent media translator 70 ("the speech processing device") to convert the received speech signals into a textual representation (Page 10: ¶[0105]: Figure 1);

"transmitting the first text message to a plurality of instant messaging devices participating in an instant messaging based conference managed by the instant messaging system" – the textual information may then be sent to a text chat interface of chat client 14, perhaps in the form of a typical chat message, via network 20 and perhaps involving IM service 22; an optional instant messaging sender 79a is depicted along connection 76 representing adaptation of the speech-to-text module 74 to carry on instant communications with chat client 14 (Page 10: ¶[0105]: Figure 1); a chat client 14 supports communications with one or more principals, and instant messaging through which text messages can be exchanged in real time with one or more other parties ("to a plurality of instant messaging devices participating in an instant messaging based conference") (Page 6: ¶[0075], Page 7: ¶[0082]);

"receiving at the speech processing device a second text message from any one among the plurality of instant messaging devices participating in the instant messaging based conference" – intelligent media translator (IMT) 70 may comprise a port for receiving textual information from a messaging client (Page 10: ¶[0104]: Figure 1); chat

client 14 may be implemented by or based upon well known instant messaging (Page 6, ¶[0075]: Figure 1);

"converting the second text message to a speech output" – intelligent media translator 70 comprises a text-to-speech conversion process for converting the received textual information into corresponding speech signals via a text-to-speech module 72 (Page 10: ¶[0103] - ¶[0104]: Figure 1);

"transmitting the speech output to a plurality of telephones participating in a teleconference managed by the teleconferencing system" – speech signals are sent through a communications medium, such as a telephone connection or RTP session, to a chat client 14 or telephone 62 (Page 10: ¶[0103] -¶[0104]: Figure 1); telephone 62 is connected through service provider 30; service provider system 30 may include, without limitation, conference call establishment (Page 7: ¶[0087]); thus, service provider system 30 is equivalent to "said teleconferencing system", which transmits teleconferencing calls through PSTN 60 to telephone 62, or to IP telephone 92 through VoIP gateway 54.

Concerning independent claim 20, *Moore et al. ('041)* discloses a speech processing device, comprising:

"convert a speech input into a text message and/or a text message into as speech output" – intelligent media translator (IMT) 70 receives speech signals, and a speech-to-text conversion process converts the received speech signals into corresponding textual information to provide the textual information ultimately to a

messaging client, receives textual information, and a text-to-speech conversion process converts the received textual information into corresponding speech signals (Page 10: ¶[0104]: Figure 1);

"receive a speech input received by said teleconferencing system from a telephone connected to the teleconferencing system" – speech information from a caller using telephone 62 is carried through PSTN 60, and is directed to speech-to-text module 74 of intelligent media translator 70 through service provider system 30 ("said teleconferencing system") via line 80 (Page 10, ¶[0118]: Figure 1); service provider system 30 may include, without limitation, conference call establishment (Page 7: ¶[0087]); thus, service provider system 30 is equivalent to "said teleconferencing system"; instant messaging (IM) service 22 communicates instant messages through chat client 14 so that text instant messages can be exchanged in real time with one or more parties (Page 7, ¶[0077]- ¶[0082]: Figure 1);

"transcribe the speech input into a first text message" – the packetized data stream is directed to speech-to-text module 74 of intelligent media translator 70 to convert the received speech signals into a textual representation (Page 10: ¶[0105]: Figure 1);

"transmit the first text message to a plurality of instant messaging devices participating in an instant messaging based conference managed by the instant messaging system" – the textual information may then be sent to a text chat interface of chat client 14, perhaps in the form of a typical chat message, via network 20 and perhaps involving IM service 22; an optional instant messaging sender 79a is depicted

along connection 76 representing adaptation of the speech-to-text module 74 to carry

on instant communications with chat client 14 (Page 10: ¶[0105]: Figure 1); a chat client

14 supports communications with one or more principals, and instant messaging

through which text messages can be exchanged in real time with one or more other

parties ("to a plurality of instant messaging devices participating in an instant messaging

based conference") (Page 6: ¶[0075], Page 7: ¶[0082]);

> "receive a second text message from any one among the plurality of instant

messaging devices participating in the instant messaging based conference" –

intelligent media translator (IMT) 70 may comprise a port for receiving textual

information from a messaging client (Page 10: ¶[0104]: Figure 1); chat client 14 may be

implemented by or based upon well known instant messaging (Page 6, ¶[0075]: Figure

1);

> "convert the second text message to a speech output" – intelligent media

translator 70 comprises a text-to-speech conversion process for converting the received

textual information into corresponding speech signals via a text-to-speech module 72

(Page 10: ¶[0103] - ¶[0104]: Figure 1);

> "transmit the speech output to a plurality of telephones participating in a

teleconference managed by the teleconferencing system" – speech signals are sent

through a communications medium, such as a telephone connection or RTP session, to

a chat client 14 or telephone 62 (Page 10: ¶[0103] -¶[0104]: Figure 1); telephone 62 is

connected through service provider 30; service provider system 30 may include, without

limitation, conference call establishment (Page 7: ¶[0087]); thus, service provider

system 30 is equivalent to "said teleconferencing system", which transmits

teleconferencing calls through PSTN 60 to telephone 62, or to IP telephone 92 through
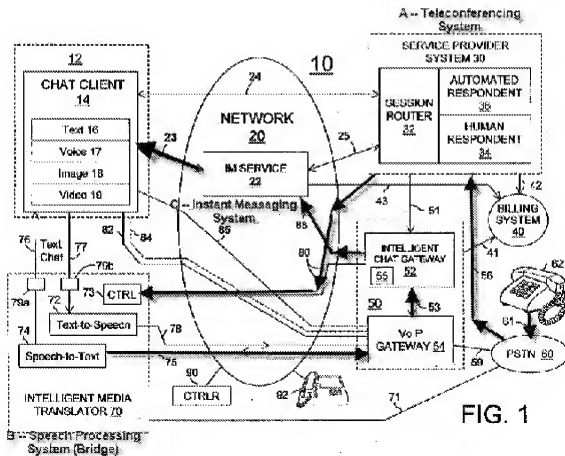
VoIP gateway 54.

  Concerning independent claims 1 and 20, *Moore et al. ('041)* further discloses:

  "wherein the speech processing device serves as a bridge between said

teleconferencing system and said messaging system, the speech processing device

being directly coupled between the teleconferencing system and the instant messaging

system and/or coupled between the teleconferencing system and the instant messaging

system via a data network, the speech processing device being configured to convert a

speech input into a text message or a text message into a speech output" – intelligent

media translator (IMT) 70 ("the speech processing device") is coupled between service

provider system 30 ("said teleconferencing system") and IM service 22 ("said

messaging system") via network 20 (Figure 1); intelligent media translator (IMT) 70 ("the

speech processing device") receives speech signals, and a speech-to-text conversion

process converts the received speech signals into corresponding textual information to

provide the textual information ultimately to a messaging client, receives textual

information, and a text-to-speech conversion process converts the received textual

information into corresponding speech signals ("being configured to convert a speech

input into a text message or a text message into a speech output") (Page 10: ¶[0104]:

Figure 1); topologically, intelligent media translator (IMT) 70 "serves as a bridge

between" service provider system 30 and IM service 22 because by following the

various arrows lines 75, 78, 80, and 86, it should be apparent that signals flow from a

telephone 62 through service provider system 30 to intelligent media translator 70, and

to IM service 22 through gateway system 50, which is basically "a data network". See

Figure 1:

FIG. 1

Concerning claims 2 and 21, *Moore et al. ('041)* discloses a profile is maintained for a given user as a preference as to how synthesized speech presented to him is rendered ("personalized voice output at the telephones"); aspects of speech rendering include whether a male or female voice is preferred, approximate speaker age, vocal characteristics, inflection, and local dialect ("a simulated voice print of the user") (¶[0113] : Figure 1).

Concerning claims 4 to 5 and 22 to 23, *Moore et al. ('041)* discloses that in the course of converting speech and other audible signals into corresponding symbols or text, IMT 70 may also perform translation among different spoken and written languages, for example, converting English text to Spanish speech and vice-versa; language preferences or compatibilities of one or both of the parties may be known or maintained in a profile database or expressed by devices ("is specified by a profile associated with said identified user"); implicitly, a user is identified in order to be associated with a profile ("identifying a user associated with said telephone"). (Page 11: ¶[0112])

Concerning claims 6 and 24, *Moore et al. ('041)* discloses that, after the packetized data stream is converted into a textual representation by speech-to-text module 74, the textual information is then sent via network 20 ("transmitting a text stream"). (Page 10: ¶[0105]: Figure 1)

Claim Rejections - 35 USC § 103

3.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

4.   Claims 1 to 2, 4 to 6, and 20 to 24 are rejected under 35 U.S.C. 103(a) as being

unpatentable over *Moore et al. ('041)*.

        Concerning independent claims 1 and 20, alternatively, the only element not

expressly disclosed by *Moore et al. ('041)* is "wherein the speech processing device

serves as a bridge between said teleconferencing system and said messaging system,

the speech processing device being directly coupled between the teleconferencing

system and the instant messaging system and/or coupled between the teleconferencing

system and the instant messaging system via a data network". *Moore et al. ('041)*

discloses all of a speech processing device being configured to convert a speech input

into a text message or a text message into speech, a teleconferencing system, and an

instant messaging system, but may connect, or couple, them in a somewhat different

way. However, it is maintained that, in an absence of unexpected advantages, it is

immaterial, and a question only of 'design choice', as to how all of the systems and

subsystems are coupled together, and what is connected to what, as it is well known

that functionalities may be distributed in an arbitrary manner within communication

networks. A mere rearrangement of parts is evidence of obviousness under legal

precedent. *In re Japikse*, 181 F.2d 1019, 86 USPQ 70 (CCPA 1950) See MPEP

§2144.04 VI. C. It would have been obvious to one having ordinary skill in the art to

rearrange the elements of the messaging response system of *Moore et al. ('041)* so that

intelligent media translator 70 "serves as a bridge between", and "is coupled between"

service provider system 30 and IM server 22 "via a data network" as a matter of "design

choice" under principles of legal precedent as a mere "rearrangement of parts" because

there would be no unexpected results obtained by a reconfiguration of elements on the

network.

Concerning claims 2 and 21, *Moore et al. ('041)* discloses a profile is maintained

for a given user as a preference as to how synthesized speech presented to him is

rendered ("personalized voice output at the telephones"); aspects of speech rendering

include whether a male or female voice is preferred, approximate speaker age, vocal

characteristics, inflection, and local dialect ("a simulated voice print of the user")

(¶[0113] : Figure 1).

Concerning claims 4 to 5 and 22 to 23, *Moore et al. ('041)* discloses that in the

course of converting speech and other audible signals into corresponding symbols or

text, IMT 70 may also perform translation among different spoken and written

languages, for example, converting English text to Spanish speech and vice-versa;

language preferences or compatibilities of one or both of the parties may be known or

maintained in a profile database or expressed by devices ("is specified by a profile

associated with said identified user"); implicitly, a user is identified in order to be

associated with a profile ("identifying a user associated with said telephone"). (Page 11:
¶[0112])

Concerning claims 6 and 24, *Moore et al. ('041)* discloses that, after the
packetized data stream is converted into a textual representation by speech-to-text
module 74, the textual information is then sent via network 20 ("transmitting a text
stream"). (Page 10: ¶[0105]: Figure 1)

Response to Arguments

5.   Applicants' arguments filed 16 September 2009 have been fully considered but
they are not persuasive.

Firstly, Applicants argue that the rejection misapprehends the burden in making
an anticipation rejection. Apparently, Applicants' position is that, although all the
elements of the claims are disclosed by *Moore et al. ('041)*, the fact that independent
claim 1 recites a method should itself be distinguishing. Applicants characterize *Moore
et al. ('041)* as disclosing "disparate and unrelated system capabilities", says the
rejection "theorizes" that they "could potentially be used together", that they are "in ways
that Moore does not even remotely suggest", and that the rejection is "clearly improper".
This position is traversed.

Applicants considerably overstate the strength of their position, and continue to
unproductively re-present the same claims without any substantial amendment. It is
maintained that the rejection over *Moore et al. ('041)* is not merely 'theoretical', and that
far from disclosing "disparate and unrelated system capabilities", *Moore et al. ('041)*

discloses what is essentially the same invention, directed to the same purposes, and with the same elements. Both Applicants' claims and *Moore et al. ('041)* are fundamentally directed to unifying instant messaging text chat and ordinary telephones through text-to-speech and speech-to-text conversion. Applicants may try to quibble about what represents an equivalent to a teleconferencing system and whether the various elements are interconnected in exactly the same way, but the basic elements of the claims are the same as *Moore et al. ('041)*.

Nor should it be required to get into a detailed analysis of one's views of the case law on inherency to show that Applicants' method claims are anticipated by *Moore et al. ('041)*. Admittedly, *Moore et al. ('041)* illustrates a system diagram in Figure 1, and not a flow chart of method steps, but the Detailed Description sets forth how each of the elements function. Thus, *Moore et al. ('041)* discloses at ¶[0103] and ¶[0104] that text-to-speech module 72 and speech-to-text module 74, respectively, perform a text-to-speech conversion for converting the received textual information into corresponding speech signals and perform a speech-to-text conversion process for converting the received speech signals into corresponding textual information. *Moore et al. ('041)* discloses, at ¶[0029] and ¶[0030], a method for managing messaging communications among a plurality of parties, and that chat accessible service system may establish communications between a chat client and another party through a computer telephony connection, a voice-enabled chat connection, or a translated text-chat-to-telephone-voice connection. There is really no need for the rejection to rely upon a doctrine of inherency; inherency might be an issue if there were any doubt as to whether one

element includes another element either by its very nature or by its definition, or

whether one element is equivalent to another element using different language. Here,

Moore et al. ('041) uses the same language to describe the same elements.

Secondly, Applicants argue that *Moore et al. ('041)* does not disclose or suggest

transmitting a speech output, converted from a text message, to a plurality of

telephones participating in a teleconference. Basically, Applicants' point here is that

they contend that "a plurality of telephones" for the chat session are not disclosed by

Moore et al. ('041). Applicants emphasis the word 'necessarily' in traversing the

rejection, thereby indicating that they think the problem is one of inherency. This is not

persuasive.

Again, it is maintained that it is not necessary to get into a theoretical discussion

of 'inherency' because the limitation of "a plurality of telephones" is expressly disclosed

by *Moore et al. ('041).* It is true that *Moore et al. ('041)* only shows two telephones in

Figure 1, *i.e.,* conventional telephone 62 and IP telephone 92. But *Moore et al. ('041)* is

plainly contemplating that each of chat client 12, conventional telephone 62, and IP

telephone 92 is representative of one of a number of similar chat clients or telephones.

Moore et al. ('041) says that chat client 14 may communicate with "one or more other

principals" although these principals "as communication appliances are not separately

depicted in FIG. 1." Similarly, *Moore et al. ('041)* says that text-to-speech module 72

and speech-to-text module 74 convert between text chat from chat client 14 and speech

signals from telephones "such as telephone 62." *Moore et al. ('041)* repeatedly

discloses that text chat communications provide for simultaneous communication

between "many different parties" and "among a plurality of parties" and "comprising a
plurality of messaging stations". (¶[0026] and ¶[0030]) A prior art reference should be
read for what it discloses to one having ordinary skill in the art. It is conventional in the
art to represent a single communications element as representative of a plurality of the
same elements. It would be manifestly unreasonable to limit *Moore et al. ('041)* to
disclosing that only one chat client can communicate at any one time with any one
telephone. This is particularly true in light of the nature of what one would commonly
understand as a 'chat room', where many participants are simultaneously
communicating together. It is contended that, if any issue exists here, it is one of
reading the prior art reference reasonably, and not any issue of inherency. Applicants'
reading limiting *Moore et al. ('041)* to only one chat client and one telephone is,
respectfully, manifestly unreasonable.

     Still, for what it's worth, it is agreed that the concept of 'inherency' includes one of
'necessity'. While not germane to any of the issues here, there can still be some doubt
as to what is involved in 'necessity'. Does the nature of 'an automobile' necessarily
include 'wheels', 'windows', 'doors', 'an engine'? It is maintained that one skilled in the
art could reasonably conclude that each of those elements is 'necessarily inherent' in
the concept of 'an automobile'. However, there could exist something that one might
call 'an automobile' that is missing its 'wheels' or missing its 'engine', *etc.* An
automobile might be stolen, stripped down, or ready to be junked, but one could
reasonably call it an automobile even if it is missing one or more of those parts.

Thirdly, Applicants argue that *Moore et al. ('041)* does not disclose transmitting a text message, transcribed from a speech output received at a speech processing device, to a plurality of instant messaging devices. Applicants point to *Moore et al. ('041)*'s disclosure of distributing to multiple operator stations 142.

Basically, Applicants are here repeating the arguments made for transmitting a speech output to now transmitting a text message, saying that there is only one single instant messaging device -- chat client 14. Again, it is maintained that this is manifestly an unreasonable reading of what is disclosed by *Moore et al. ('041)*. Of course, Applicants' arguments are responsive insofar as they bring up new points to dispute, but it is not understood how Applicants can reasonably doubt that a communication among a plurality of chat clients and a plurality of telephones through text-to-speech and speech-to-text media translation is disclosed by *Moore et al. ('041)* Applicants admit that the rejection 'repeatedly' points to disclosure supporting multiple communications among multiple parties – by *Moore et al. ('041)* at ¶[0030], ¶[0075], and ¶[0083]. Applicants' citation of the disclosure by *Moore et al. ('041)* of multiple operator stations 142 does not appear to be relevant, and was not relied upon in the rejection.

Finally, Applicants address the rejection of independent claims 1 and 20 under 35 U.S.C. 103(a) as being unpatentable over *Moore et al. ('041)*. Applicants note the citation of *In re Japikse*, 181 F.2d 1019, 86 USPQ 70 (CCPA 1950) and MPEP §2144.04 VI. C. However, Applicants discuss the facts of the position of the starting switch in *Japikse*, and argue that these facts are entirely unrelated to the facts here of bridging a teleconferencing system and a messaging system. Moreover, Applicants

note that *Japikse* requires a motivation or reason should be provided to make the proposed rearrangement of parts.

It is maintained that any rearrangement of parts in a communication network may be an obvious matter of 'design choice' in some instances. The facts here are quite similar to those in *Japikse*. Rearrangement of a position of a switch would not change the overall function of the device, just as a rearrangement of the relative position of the speech processing device would not affect its function. The rationale is as follows: Although some devices on a communication network could be required to execute functions before other devices on a communication network, it is common for devices on a communication network to be distributed so that essentially any device is accessible to any other device. There would not appear to be any substantial advantage to positioning the speech processing device in any particular location with respect to an instant messaging system and a teleconferencing system because all of the components are likely to be located at geographically remote positions in the network. Applicants' Specification does not provide any advantage of positioning the speech processing device in any specific location with regard to the instant messaging system and the teleconferencing system. One skilled in the art would recognize that there would be no substantial difference in function as long as all the devices and systems are accessible through the communication network to every other device and system because they are, in any event, likely to be placed at remotely distributed locations.

Therefore, a rejection of claims 1 to 2, 4 to 6, and 20 to 24 under 35 U.S.C.

§102(e) as being anticipated by *Moore et al. ('041)*, and/or a rejection of claims 1 to 2, 4

to 6, and 20 to 24 under 35 U.S.C. §103(a) as being obvious over *Moore et al. ('041)*,

are proper.

Conclusion

6.   **THIS ACTION IS MADE FINAL.** Applicants are reminded of the extension of

time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Martin Lerner whose telephone number is (571) 272-

7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to

Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone

number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Martin Lerner/
Primary Examiner
Art Unit 2626
November 2, 2009